

1. In this question, we will give unequal weightage to errors in linear regression setup.

For  $i^{th}$  observation, the error is given by  $e_i = (y_i - \hat{y}_i)$ , where  $y_i$  is the actual value and  $\hat{y}_i$  is the predicted value. We weigh the error by  $r_i$ . Our objective is to minimize the weighted sum of squared errors, i.e.,  $\sum_{i=1}^n r_i e_i^2$ . Assuming you can form a vector/matrix from  $r_i$ 's, write down the objective function in matrix form, and find the optimal  $\theta$ 's.

2. Given a simple model,  $\hat{y} = \theta_1 x$ , where  $x$  is the input and  $\theta_1$  is the parameter to be estimated. What is the optimal value of  $\theta_1$  in terms of the data?
3. Given a simple model,  $\hat{y} = \theta_0$ , where  $x$  is the input and  $\theta_0$  is the parameter to be estimated. What is the optimal value of  $\theta_0$  in terms of the data?
4. Suppose I want to put a hard constraint for linear regression that all the coefficients should be positive. How can I do that?
5. Prove that stochastic gradient is an unbiased estimator of the true gradient.
6. Prove that the expected size of bootstrapped sample 0.632 times the original sample size.